

nucleic acid obtained from said non-cancerous biological sample; generating a matrix of sequencing tag counts for said at least one promoter based on said mapping; analysing said matrix of sequencing tag counts; and determining the differential enrichment of the at least one promoter in the nucleic acid relative to the at least one promoter in the reference nucleic acid using the analysis of said matrix of sequencing tag counts, wherein the differential enrichment of the at least one promoter in the cancerous biological sample obtained from the subject relative to that in the non-cancerous sample is indicative of the presence of a promoter associated with cancer in a subject.

[0010] In another aspect there is provided a method for determining the activity of at least one promoter in a cancerous biological sample relative to a non-cancerous biological sample, comprising; mapping isolated nucleic acid comprising at least one promoter sequence obtained from said cancerous biological sample against a reference nucleic acid obtained from said non-cancerous biological sample; generating a matrix of sequencing tag counts for said at least one promoter based on said mapping; analysing said matrix of sequencing tag counts; and determining the differential activity of the at least one promoter in the nucleic acid relative to the at least one promoter in the reference nucleic acid using the analysis of said matrix of sequencing tag counts.

[0010a] In another aspect, there is provided the use of a biomarker comprising at least one promoter having increased activity in a cancerous biological sample relative to a normal non-cancerous biological sample, wherein the promoter comprises an increase of SUZ12 binding sites relative to the total promoter population in the manufacture of a medicament for detecting cancer in a subject.

[0010b] In another aspect, there is provided a method for the determining the prognosis of a subject with cancer comprising,

mapping an isolated nucleic acid comprising at least one promoter obtained from a cancerous biological sample of the subject against a reference nucleic acid obtained from a non-cancerous biological sample to obtain an RPKM or FPKM value for said at least one promoter; and

determining the differential activity of the at least one promoter in the nucleic acid relative to the activity of the at least one promoter in the reference nucleic acid using said RPKM or FPKM value, wherein an increased activity of the at least one promoter in the

cancerous sample relative to that in the non-cancerous sample is indicative of the susceptibility of the subject to cancer.

[0010c] In another aspect, there is provided a method for determining the presence of at least one promoter associated with cancer in a cancerous biological sample relative to a non-cancerous biological sample, comprising;

mapping isolated nucleic acid comprising at least one promoter sequence obtained from said cancerous biological sample against a reference nucleic acid obtained from said non-cancerous biological sample to obtain a read per kilo-base per million (RPKM) value or fragments per kilo-base per million (FPKM) value for said at least one promoter;

generating a matrix of sequencing tag counts for said at least one promoter based on said mapping;

analysing said matrix of sequencing tag counts; and

determining the differential enrichment of the at least one promoter in the nucleic acid relative to the at least one promoter in the reference nucleic acid using the analysis of said matrix of sequencing tag counts, wherein the differential enrichment of the at least one promoter in the cancerous biological sample obtained from the subject relative to that in the non-cancerous sample is indicative of the presence of a promoter associated with cancer in a subject,

wherein the nucleic acid is isolated from said cancerous biological sample by immunoprecipitation of chromatin achieved by an antibody specific for a modified histone protein; and

wherein the modified histone protein comprises at least one histone modification selected from the group consisting of H3K4me3, H3K4me1 and H3K27ac.

[0010d] In another aspect, there is provided a method for determining the activity of at least one promoter in a cancerous biological sample relative to a non-cancerous biological sample, comprising;

mapping isolated nucleic acid comprising at least one promoter sequence obtained from said cancerous biological sample against a reference nucleic acid obtained from said non-cancerous biological sample to obtain a read per kilo-base per million (RPKM) value or fragments per kilo-base per million (FPKM) value for said at least one promoter;

generating a matrix of sequencing tag counts for said at least one promoter based on said mapping;

analysing said matrix of sequencing tag counts; and

determining the differential activity of the at least one promoter in the nucleic acid relative to the at least one promoter in the reference nucleic acid using the analysis of said matrix of sequencing tag counts,

5 wherein the nucleic acid is isolated from said cancerous biological sample by immunoprecipitation of chromatin achieved by an antibody specific for a modified histone protein; and

 wherein the modified histone protein comprises at least one histone modification selected from the group consisting of H3K4me3, H3K4me1 and H3K27ac.

10

DEFINITIONS

[0011] The term “antigen binding protein” as used herein refers to antibodies, antibody fragments and other protein constructs, such as domains, which are capable of binding to an antigen.

15 [0012] The term “antibody” is used herein in the broadest sense to refer to molecules with an immunoglobulin-like domain and includes monoclonal, recombinant, polyclonal, chimeric, humanised, bispecific and heteroconjugate antibodies; a single variable domain, a domain antibody, antigen binding fragments, immunologically effective fragments, single chain Fv, diabodies, Tandabs™, etc (for a summary of alternative “antibody” formats see Holliger and Hudson, Nature Biotechnology, 2005, Vol 23, No. 9, 1126-1136).

20 [0013] The phrase “single variable domain” refers to an antigen binding protein variable domain (for example, V_H, V_{HH}, V_L) that specifically binds an antigen or epitope independently of a different variable region or domain.

Claims

1. A method for determining the activity of at least one promoter in a cancerous biological sample relative to a non-cancerous biological sample, comprising;
 - 5 mapping isolated nucleic acid comprising at least one promoter sequence obtained from said cancerous biological sample against a reference nucleic acid obtained from said non-cancerous biological sample to obtain a read per kilo-base per million (RPKM) value or fragments per kilo-base per million (FPKM) value for said at least one promoter; and
 - 10 determining the differential activity of the at least one promoter sequence in the nucleic acid relative to the activity of the at least one promoter in the reference nucleic acid sequence using said RPKM or FPKM value.
2. The method of claim 1, wherein the cancerous and non-cancerous biological sample
 - 15 comprises a single cell, multiple cells, fragments of cells, body fluid or tissue.
3. The method of any one of claims 1-2, wherein the cancerous and non-cancerous biological sample is obtained from the same subject.
- 20 4. The method of any one of claims 1 or 2, wherein the cancerous and non-cancerous biological sample are each obtained from different subjects.
5. The method of any one of claims 1-4, wherein the nucleic acid is isolated from said cancerous biological sample by immunoprecipitation of chromatin, wherein said
 - 25 nucleic acid comprises said at least one promoter.
6. The method of claim 5, wherein immunoprecipitation of chromatin is achieved by an antibody specific for a modified histone protein.
- 30 7. The method of claim 6, wherein the modified histone protein comprises at least one histone modification selected from the group consisting of H3K4me3, H3K4me1 and H3K27ac.

8. The method of claim 7, wherein the antibody is specific to the at least one histone modification selected from the group consisting of H3K4me3, H3K4me1 and H3K27ac.

5

9. The method of any one of claims 1-8, wherein the isolated nucleic acid comprising at least one promoter is amplified with at least one primer.

10

10. The method of claim 9, wherein said amplified nucleic acid is used to construct a nucleic acid sequence library with said amplified nucleic acid.

15

11. The method of any one of claims 1-10, wherein the mapping step comprises calculating the RPKM values based upon the total sequence tags for the at least one promoter in the mapped nucleic acid relative to the reference nucleic acid.

20

12. The method of any one of claims 1-10, wherein the mapping step comprises calculating the FPKM values based upon identified transcript sequences associated with the at least one promoter in the mapped nucleic acid relative to the reference nucleic acid.

25

13. The method of any one of claims 1-12, wherein the step of determining the differential activity of the at least one promoter sequence comprises determining that the RPKM or FPKM value for the at least one promoter in the nucleic acid obtained from the cancerous biological sample is:

30

- i) greater than between a 1 to 20-fold, such as a 1-fold, 2-fold, 3-fold, 4-fold or 5-fold, change in mean RPKM or FPKM value relative to the RPKM or FPKM value of the at least one promoter in the reference nucleic acid obtained from the non-cancerous biological sample; and
- ii) greater than a 0.1 RPKM or FPKM range relative to the RPKM or FPKM value of the at least one promoter in the reference nucleic acid obtained from the non-cancerous biological sample.

14. The method of any one of claims 1-13, wherein the at least one promoter comprises an increase of SUZ12 binding sites relative to the total promoter population.
- 5 15. The method of any one of claims 1-14, wherein the at least one promoter is positioned adjacent to a gene associated with cell-type specification, embryonic development or transcription factors.
16. The method of claim 15, wherein the at least one promoter is positioned adjacent to a gene associated with cancer.
- 10 17. The method of claim 16, wherein the gene is NKX6-3, SALL4, HOXB9, MET, TNK2, KLK1, FAR2, HOXA11 or HOXA11-AS.
18. The method of any of claims 1-17, wherein the cancer is gastric cancer.
- 15 19. The method of any of claims 1-18, wherein the at least one promoter comprises a cryptic promoter.
- 20 20. A method for the determining the prognosis of a subject with cancer comprising,
mapping an isolated nucleic acid comprising at least one promoter obtained
from a cancerous biological sample of the subject against a reference nucleic acid
obtained from a non-cancerous biological sample to obtain an RPKM or FPKM value
for said at least one promoter; and
determining the differential activity of the at least one promoter in the nucleic
25 acid relative to the activity of the at least one promoter in the reference nucleic acid
using said RPKM or FPKM value, wherein an increased activity of the at least one
promoter in the cancerous sample relative to that in the non-cancerous sample is
indicative of the susceptibility of the subject to cancer.
- 30 21. A method for determining the presence of at least one promoter associated with cancer in a subject, comprising:

mapping isolated nucleic acid comprising at least one promoter obtained from a cancerous biological sample of the subject against a reference nucleic acid obtained from a non-cancerous biological sample to obtain an RPKM or FPKM value for said at least one promoter; and

5 determining the differential activity of the at least one promoter in the nucleic acid relative to the activity of the at least one promoter in the reference nucleic acid using said RPKM or FPKM value, wherein an increased activity of the at least one promoter in the cancerous biological sample obtained from the subject relative to that in the non-cancerous sample is indicative of the presence of a promoter associated
10 with cancer in a subject.

22. The method of claim 21, wherein the at least one promoter associated with cancer is present when an RPKM or FPKM value for the at least one promoter in the nucleic acid obtained from the biological sample is:

15 i) greater than between a 1 to 20-fold, such as a 1-fold, 2-fold, 3-fold, 4-fold or 5-fold, change in mean RPKM or FPKM value relative to the RPKM or FPKM value of the at least one promoter in the reference nucleic acid obtained from the non-cancerous biological sample; and

20 ii) greater than a 0.1 RPKM or FPKM range relative to the RPKM or FPKM value of the at least one promoter in the reference nucleic acid obtained from the non-cancerous biological sample.

23. The method of claim 21, wherein the at least one promoter comprises an increase in SUZ12 binding sites relative to the total promoter population.

25 24. The method of claim 23, wherein the at least one promoter exhibits a low DNA methylation level relative to the total promoter population.

30 25. The method of any one of claim 23 or 24, wherein the at least one promoter is positioned adjacent to a gene associated with cell-type specification, embryonic development or transcription factors.

26. The method of claim 25, wherein the gene is NKX6-3, SALL4, HOXB9, MET, TNK2, KLK1, FAR2, HOXA11 or HOXA11-AS.

27. The method of any one of claims 23-26, wherein the at least one promoter comprises
5 a cryptic promoter.

28. A method for determining the presence of at least one promoter associated with cancer in a cancerous biological sample relative to a non-cancerous biological sample, comprising;

10 mapping isolated nucleic acid comprising at least one promoter sequence obtained from said cancerous biological sample against a reference nucleic acid obtained from said non-cancerous biological sample to obtain a read per kilo-base per million (RPKM) value or fragments per kilo-base per million (FPKM) value for said at least one promoter;

15 generating a matrix of sequencing tag counts for said at least one promoter based on said mapping;

analysing said matrix of sequencing tag counts; and

20 determining the differential enrichment of the at least one promoter in the nucleic acid relative to the at least one promoter in the reference nucleic acid using the analysis of said matrix of sequencing tag counts, wherein the differential enrichment of the at least one promoter in the cancerous biological sample obtained from the subject relative to that in the non-cancerous sample is indicative of the presence of a promoter associated with cancer in a subject,

25 wherein the nucleic acid is isolated from said cancerous biological sample by immunoprecipitation of chromatin achieved by an antibody specific for a modified histone protein; and

wherein the modified histone protein comprises at least one histone modification selected from the group consisting of H3K4me3, H3K4me1 and H3K27ac.

29. A method for determining the activity of at least one promoter in a cancerous biological sample relative to a non-cancerous biological sample, comprising;

mapping isolated nucleic acid comprising at least one promoter sequence obtained from said cancerous biological sample against a reference nucleic acid obtained from said non-cancerous biological sample to obtain a read per kilo-base per million (RPKM) value or fragments per kilo-base per million (FPKM) value for said at least one promoter;

generating a matrix of sequencing tag counts for said at least one promoter based on said mapping;

analysing said matrix of sequencing tag counts; and

determining the differential activity of the at least one promoter in the nucleic acid relative to the at least one promoter in the reference nucleic acid using the analysis of said matrix of sequencing tag counts,

wherein the nucleic acid is isolated from said cancerous biological sample by immunoprecipitation of chromatin achieved by an antibody specific for a modified histone protein; and

wherein the modified histone protein comprises at least one histone modification selected from the group consisting of H3K4me3, H3K4me1 and H3K27ac.

30. The method of claim 28 or 29, wherein the generating step comprises calculating the matrix based upon a sequence tag count for the at least one promoter in the mapped nucleic acid relative to the reference nucleic acid.

31. The method of claim 28 or 29, wherein the analysis step comprises analyzing the matrix using a DESeq2 algorithm.

32. The method of claim 28 or 29, wherein the at least one promoter is positioned adjacent to a gene associated with cancer.

33. The method of claim 32, wherein the gene is RASA3, GRIN2D, TNNI3, SHD, ATP10B, SMTN, MYO15B, C2orf61, LINC00443 or ACHE.

34. The method of claim 28, wherein the differential enrichment is identified based upon a FDR rate of 10% and an absolute fold change of greater than 1.5.
 35. The method of claim 29, wherein the differential activity is identified based upon a FDR rate of 10% and an absolute fold change of greater than 1.5.
- 5